JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 48, 155–164 (1974)

Descent Methods in Smooth, Rotund Spaces with Applications to Approximation in L^p

FRANCIS E. SULLIVAN* AND BARBRO ATLESTAM

Chalmers tekniska högskola, Institutionen för Informationsbehandling, Fack, 402 20 Göteborg, Sweden

Submitted by Germund Dahlquist

1. INTRODUCTION

The purpose of this paper is to adapt some of the ideas used for approximation on finite-dimensional subspaces of L^2 to the more general setting of smooth, rotund spaces. A general principle for construction of approximation methods is obtained. As an application we give a unified treatment of some of the L^p-approximation methods found in the literature, both for, $1 < p \leqslant 2$ and $2 \leqslant p < \infty$. By employing sequences of L^{p_i} approximations with $p_i \to 1$ or $p_i \to \infty$ these methods can be used for L^1 or L^∞ approximations.

Section 2 contains background material on smoothness, rotundity, and orthogonality along with presentation of some properties of the best approximation operator.

One of the L^2 ideas which can be generalized is the Gram–Schmidt orthogonalization method. An imitation of the Gram–Schmidt process results in the method of "alternation" with B-operators which is given in Section 3.

Another property of approximation in L^2 is that the best approximation operator is given by a contractive linear projection. In Section 4 we obtain a theorem which gives techniques for obtaining best approximations using sequences of linear operators. These are based on using the derivatives of the norm functional to define B-operators. One of these involves the second derivative of the norm-functional and is a sequence of "weighted" L^2-projections. We show the connection between this and the derivative of the best approximation operator. We also consider "mixed" methods which combine alternation with sequences of linear operators. In Section 5 we show how the methods apply to L^p and comment briefly on some numerical experiments.

* Present address: Department of Mathematics, The Catholic University of America, Washington, D.C. 20017.

2. Background

Let X be a real Banach space with unit sphere S and X^* the dual with unit sphere S^*. For each $f \in X, f \neq 0$ there exists, by the Hahn–Banach theorem, a $g^* \in S^*$ with $g^*(f) = \|f\|$. We call g^* a normer for f. Smoothness is the statement that each $f \in X$, $f \neq 0$ has but one normer denote by $n(f)$, and

$$n(f)(h) = \lim_{\lambda \to 0}(\|f + \lambda h\| - \|f\|)/\lambda \qquad \text{for all } h \in X.$$

The map $f \to n(f)$ is continuous away from 0 [2, p. 149].

James [4] defines orthogonality, $f \overset{\perp}{\to} g$, iff $\|f + \lambda g\| \geqslant \|f\|$ for all real λ. In smooth spaces this is equivalent to $n(f)(g) = 0$. If M is a closed subspace of X then a vector $\bar{m} \in M$ minimizes $\|f - m\|$, $m \in M$ iff $f - \bar{m} \overset{\perp}{\to} m$ for all $m \in M$. If M is finite-dimensional with basis $m_1, m_2, \ldots, m_k$ and X is smooth, this can be expressed by

$$n(f - \bar{m})(m_i) = 0, \qquad 1 \leqslant i \leqslant k.$$

The space X is rotund if each $g^* \in S^*$ norms at most one $f \in S$. Thus if $\|f\| = 1$ and X is smooth and rotund $n(f)(g) = \|g\|$ iff $f = g$. Rotundity can also be expressed by the geometrical condition of strict convexity, i.e., $\|f + g\| = \|f\| + \|g\|$ iff f and g are linearly dependent. This means that if $M \subseteq X$ is a closed subspace and $f \in X$, the best approximation to f in M is unique when it exists.

If X is smooth, rotound and M is finite-dimensional then we have a continuous (nonlinear) operator $P_M: X \to M$ with $P_M(f)$ the best approximation to f in M. This operator is characterized by the fact that $f - P_M(f) \overset{\perp}{\to} M$ for all $f \in X$. Note that for $m \in M$, $P_M(f + m) = P_M(f) + m$.

3. The Alternation Method

Unless explicity stated otherwise, we assume that X is smooth, rotund and that M is a finite-dimensional subspace of X. Generalizing the properties of P_M slightly we say that a continuous map $P: X \backslash M \to M$ has property B relative to M if:

(1) $\|f - P(f)\| \leqslant \|f\|$ for all $f \in X \backslash M$

(2) $P(f) = 0$ iff $f \overset{\perp}{\to} M$ iff $\|f - P(f)\| = \|f\|$.

We then have the following analog of the Gram–Schmidt process.

THEOREM 1. *Let M be a finite-dimensional subspace with*

$$M = M_1 \oplus M_2 \cdots \oplus M_n$$

and for each i, P_i a continuous map $X \backslash M_i \to M_i$ with property B relative to M_i. Define the map A on $X \backslash M$ by $A(h) = g_n$ where $g_0 = h$ and $g_{i+1} = g_i - P_{i+1}(g_i)$. Then for each $f \in X \backslash M$, $\lim_{k\to\infty} A^k(f) = f - P_M(f)$. Here $A^1(f) = A(f)$ and $A^{k+1}(f) = A(A^k(f))$.

Proof. First note for $f \in X\backslash M$, $m \in M$, $\| A(f-m)\| = \| f - m \|$ iff $m = P_M(f)$. Also, for $h \in X \backslash M$ arbitrary $\| A(h)\| \leqslant \| h \|$. Thus $\| A^k(f)\|$ is decreasing and since M is finite-dimensional there is a convergent subsequence $A^{k_j}(f) \to g$. We shall show that $g = f - P_M(f)$.

From the form of A we have that $f - A^{k_j}(f) \in M$ for each k_j and so $f - g \in M$. If $f - g \neq P_M(f)$ then $\| A(g)\| < \| g \|$. However, by continuity and the fact that $\| A^k(f)\|$ converges

$$\| A(g)\| = \| A(\lim A^{k_j}(f))\| = \lim \| A^{k_j}(f)\| = \| g \| .$$

Thus $g = f - P_M(f)$ and since every convergent subsequence has the same limit $A^k(f)$ converges to $f - P_M(f)$. Q.E.D.

Note that if the P_i are, in fact, the best approximation operators on the subspaces M_i we have the theorem of Stiles [7]. The case of two disjoint subspaces may be presented in a new corollary.

COROLLARY. *Let M_1 and M_2 be disjoint subspaces. Then*

$$P_{M_1 \oplus M_2}(f) = m + P_{M_2}(f - m),$$

where

$$m = P_{M_1}(f - P_{M_2}(f - m)).$$

Proof. Define the sequence m_i, $i = 1, 2,\ldots$, by

$$m_1 = P_{M_1}(f), \qquad m_{i+1} = P_{M_1}(f - P_{M_2}(f - m_i)), \quad i = 1, 2, \ldots.$$

The mapping A of Theorem 1 takes the form

$$A(f) = f - P_{M_1}(f) - P_{M_2}(f - P_{M_1}(f)).$$

Using induction on i we establish

$$A^i(f) = f - m_i - P_{M_2}(f - m_i).$$

Putting $m = \lim m_i$ when $i \to \infty$ we find, using Theorem 1,

$$f - P_{M_1 \oplus M_2}(f) = f - m - P_{M_2}(f - m)$$

from which the desired result follows. Q.E.D.

This result is obtained by Holmes and Kripke [3] by a different argument.

The technique of the above theorem does not appear to apply to the case in which M is not finite-dimensional. However, if we assume that X is *uniformly* rotund we can obtain best approximations in subspaces defined by increasing unions of finite-dimensional spaces from limits of sequences given by B-operators.

X is called uniformly rotund if for

$$\|f\| = \|g\| = 1 \qquad \text{and} \qquad \|(f-g)/2\| \geqslant \epsilon > 0$$

there exist a $\delta(\epsilon) > 0$ such that

$$\|(f+g)/2\| \leqslant 1 - \delta(\epsilon).$$

A uniformly rotund space is reflexive.

Lemma 1. *If X is uniformly rotund and $\{M_i\}$ is an increasing sequence of closed subspaces with $\overline{M} \equiv \mathrm{cl}(\cup M_i)$, then for all $f \in X \backslash \overline{M}$ the sequence $\{P_{M_i} f\}$ converges to $P_{\overline{M}} f$.*

Proof. First note that the sequence of reals $\|f - P_{M_i} f\|$ is bounded decreasing and hence convergent. If $\{P_{M_i} f\}$ is not convergent there is a subsequence and an $\epsilon' > 0$ such that $\| P_{M_i} f - P_{M_j} f \| > \epsilon'$ for all $i \neq j$ in the subsequence. Hence, since $\|f - P_{M_i} f\|$ converges there is an $\epsilon > 0$ with

$$\frac{1}{2} \left\| \frac{f - P_{M_i} f}{\|f - P_{M_i} f\|} - \frac{f - P_{M_j} f}{\|f - P_{M_j} f\|} \right\| > \epsilon.$$

Thus, from uniform rotundity there is a $\delta(\epsilon) > 0$ with

$$\frac{1}{2} \left\| \frac{f - P_{M_i} f}{\|f - P_{M_i} f\|} + \frac{f - P_{M_j} f}{\|f - P_{M_j} f\|} \right\| \leqslant 1 - \delta(\epsilon).$$

If we assume that $i > j$ then using the contraction property of the best approximation operator we have

$$\frac{1}{2} \left(\frac{1}{\|f - P_{M_i} f\|} + \frac{1}{\|f - P_{M_j} f\|} \right) \|f - P_{M_i} f\| \leqslant 1 - \delta(\epsilon).$$

Again using the convergence of $\| f - P_{M_i} f \|$ we obtain a contradiction and so $\{P_{M_i} f\}$ converges.

Assume that $\lim P_{M_i} f = \bar{m}$.

Let $m \in \overline{M}$ and $\epsilon > 0$ be given. Then there is an i and an $m' \in M_i$ such that

$$\| f - \bar{m} + m \| \geqslant \| f - \bar{m} + m' \| - \epsilon/2.$$

Then we can find $j_0 > i$ such that

$$\| f - \bar{m} + m' \| \geqslant \| f - P_{M_j} f + m' \| - \epsilon/2, \qquad j \geqslant j_0$$

giving

$$\| f - \bar{m} + m' \| \geqslant \| f - P_{M_j} f \| - \epsilon/2.$$

Hence

$$\| f - \bar{m} + m \| \geqslant \| f - P_{M_j} f \| - \epsilon.$$

Letting $j \to \infty$ we obtain

$$\| f - \bar{m} + m \| \geqslant \| f - \bar{m} \| - \epsilon.$$

Since ϵ is arbitrary we must conclude that $f - \bar{m} \xrightarrow{\perp} \overline{M}$. Q.E.D.

Given B-operators P and Q with ranges M and N, respectively, we say that P is less than Q if $M \subseteq N$ and for all $f \in X \backslash N$ and all $n \geqslant 0$

$$\|(I - Q)^n f \| \leqslant \|(I - P)^n f \|.$$

THEOREM 2. *Let X be uniformly rotund and P_i be an increasing sequence of B-operators with finite-dimensional ranges M_i. Then if $\overline{M} \equiv \mathrm{cl}(\cup M_i)$ for any $f \in X \backslash M$ the sequence $f - m_i$ where $m_0 = 0$ and*

$$f - m_{i+1} = f - m_i - P_{i+1}(f - m_i)$$

converges to $f - P_{\overline{M}} f$.

Proof. Since the P_i are increasing we have that for all j,

$$\| f - m_{i+j} \| = \|(I - P_{i+j-1}) \cdots (I - P_i)(f - m_i)\| \leqslant \|(I - P_i)^j (f - m_i)\|.$$

From Theorem 1 follows that for i fixed

$$(I - P_i)^j (f - m_i) \to f - P_{M_i} f \qquad \text{as} \qquad j \to \infty.$$

Using Lemma 1 shows that $\| f - m_i \|$ converges to $\| f - P_{\overline{M}} f \|$. Imitating the argument of the first part of the proof of Lemma 1 we have that m_i converges. Then $f - m_i$ converges to $f - P_{\overline{M}} f$. Q.E.D.

4. Construction of B-Operators

We now consider some methods for defining the maps P_i of Theorem 1. We assume that M is finite-dimensional. Since n is a map from X to X^*, for each $g \in X$, $g \neq 0$, $n'(g)$, where n' is the Fréchet derivative of n, is a bilinear form on X when it exists. If $\Psi\colon R^+ \to R^+$ is a strictly increasing twice continuously differentiable convex function then $\phi(f) \equiv \Psi(\| f \|)$ also has two derivatives where from the chain rule

$$\begin{aligned} \phi'(f)(g) &= \Psi'(\| f \|)\, n(f)(g), \\ \phi''(f)(g, h) &= \Psi'(\| f \|)\, n'(f)(g, h) + \Psi''(\| f \|)\, n(f)(g)\, n(f)(h). \end{aligned} \tag{*}$$

Note that minimizing $\| f - m \|$ is equivalent to minimizing $\phi(f - m)$ and $f - \bar{m} \xrightarrow{\perp} M$ iff $\phi'(f - \bar{m})(m) = 0$ for all $m \in M$. We say that M has property T for ϕ if for each $g \in X\backslash M$ there exist nonzero constants c_1, c_2 such that

$$c_1 \| h \|^2 \geqslant \phi''(g)(h, h) \geqslant c_2 \| h \|^2 \qquad \text{for all } h \in M.$$

In Section 5 we give examples of functions Ψ and subspaces of L^p, $p \geqslant 2$ having this property.

Theorem 1 requires continuity of the operators P_i. This is based on the following lemma which is of some independent interest.

Lemma 2. *If $p\colon X\backslash M \to M$ is a continuous operator such that $p(f) = 0$ iff $f \xrightarrow{\perp} M$ then the map $f \to \alpha p(f)$, where α minimizes $\| f - \alpha p(f)\|$ is also continuous.*

Proof. Suppose we have a sequence $f_i \to f$. From the assumed continuity $p(f_i) \to p(f)$. We shall show that $\alpha_i p(f_i) \to \alpha p(f)$ where α minimizes $\| f - \alpha p(f)\|$. Note that for each i, $f_i - \alpha_i p(f_i) \xrightarrow{\perp} \alpha_i p(f_i)$ and so

$$\| \alpha_i p(f_i)\| \leqslant 2 \| f_i \| .$$

If $p(f) \neq 0$ then since $| \alpha_i | \leqslant 2 \| f_i \| / \| p(f_i)\|$, by passing to a subsequence we may assume $\alpha_i \to \bar{\alpha}$. If $\| f - \bar{\alpha} p(f)\|$ is not minimized then from uniqueness of best approximation for some $\delta > 0$ we have

$$\delta \leqslant | n(f - \bar{\alpha} p(f))(p(f))| .$$

However since n is continuous for each $\epsilon > 0$ and i large we have

$$\begin{aligned} | n(f - \bar{\alpha} p(f))(p(f))| &\leqslant | n(f_i - \alpha_i p(f_i))(p(f))| + \epsilon \\ &= | n(f_i - \alpha_i p(f_i))(p(f) - p(f_i))| + \epsilon \\ &\leqslant \| p(f) - p(f_i)\| + \epsilon. \end{aligned}$$

Since $p(f_i) \to p(f)$ this gives $\epsilon \geqslant \delta$ for all $\epsilon > 0$, a contradiction.

If $p(f) = 0$ then $f \xrightarrow{\perp} M$ and we shall show that $\alpha_i p(f_i) \to 0$. From the boundedness of $\alpha_i p(f_i)$ and passing to a subsequence we may assume $\alpha_i p(f_i) \to m$. For $\epsilon > 0$ and i large we have

$$\|f\| \geqslant \|f_i\| - \epsilon/2 \qquad \text{and} \qquad \|f_i - \alpha_i p(f_i)\| \geqslant \|f - m\| - \epsilon/2.$$

Since $f \xrightarrow{\perp} M$ we have $\|f - m\| \geqslant \|f\|$.

Combining these gives

$$\|f_i\| \geqslant \|f_i - \alpha_i p(f_i)\| \geqslant \|f - m\| - \epsilon/2 \geqslant \|f\| - \epsilon/2 \geqslant \|f_i\| - \epsilon.$$

In the limit we have $\|f\| = \|f - m\|$ and so $m = 0$ by uniqueness of best approximation. Q.E.D.

THEOREM 3. *Let M have basis $\{m_1, m_2, \ldots, m_k\}$ and let ϕ be a map of the type described above. Then the operator $P_1\colon X\backslash M \to M$ given by $P_1(f) = \alpha_1 p_1(f)$ where $p_1(f) = \sum_{i=1}^k \phi'(f)(m_i)\, m_i$ and α_1 is chosen to minimize $\|f - \alpha_1 p_1(f)\|$ is continuous and has property B.*

If in addition, for each fixed $m \in M$ the map $X\backslash M \to X^*$ where $f \to \phi''(f)(m, \cdot)$ is defined and continuous and M has property T for ϕ then $P_2(f) = \alpha_2 p_2(f)$, $X\backslash M \to M$, where $p_2(f)$ satisfies

$$\sum_{i=1}^{k} \phi''(f)(m_i, p_2(f))\, m_i = p_1(f)$$

and α_2 is chosen to minimize $\|f - \alpha_2 p_2(f)\|$, also is continuous and has property B.

Proof. For P_1, $\|f - P_1(f)\| \leqslant \|f\|$ with equality if $P_1(f) = 0$ is clear. If $f \xrightarrow{\perp} M$ then $p_1(f) = 0$ so $P_1(f) = 0$. If $P_1(f) = 0$ then either $\alpha_1 = 0$ or $p_1(f) = 0$. When $\alpha_1 = 0$ we have $f \xrightarrow{\perp} p_1(f)$ so from equation (*)

$$\sum_{i=1}^{k} \Psi'(\|f\|)\,(n(f)(m_i))^2 = 0$$

and $f \xrightarrow{\perp} M$. Otherwise $p_1(f) = 0$ and $f \xrightarrow{\perp} M$ is clear. If we have $\|f - P_1(f)\| = \|f\|$ then $f \xrightarrow{\perp} p_1(f)$ and so $f \xrightarrow{\perp} M$ follows in the same way. Continuity follows from Lemma 2 and the continuity of n and Ψ.

In the case P_2, condition T for ϕ on M says that

$$h \to \sum_{i=1}^{k} \phi''(f)(m_i, h)\, m_i$$

is a bounded linear operator from M to M with bounded inverse. Hence $p_2(f)$ is defined. The condition $\|f - P_2(f)\| \leqslant \|f\|$ with equality if $P_2(f) = 0$ is again clear. If $f \xrightarrow{\perp} M$ then $p_1(f) = 0$ so $p_2(f) = 0$. If $P_2(f) = 0$ then either $\alpha_2 = 0$ or $p_2(f) = 0$. In case $\alpha_2 = 0$ we have $f \xrightarrow{\perp} p_2(f)$. Let

$$p_2(f) = \sum_{i=1}^{k} \lambda_i m_i$$

then we have that

$$\phi''(f)(m_i, p_2(f)) = \phi'(f)(m_i)$$

for each i and

$$\begin{aligned} \phi''(f)(p_2(f), p_2(f)) &= \sum \lambda_i \phi''(f)(m_i, p_2(f)) = \sum \lambda_i \phi'(f)(m_i) \\ &= \phi'(f)(p_2(f)) = 0. \end{aligned}$$

From property T we have $p_2(f) = 0$ then from property T again, $p_1(f) = 0$ and $f \xrightarrow{\perp} M$. If $p_2(f) = 0$ then from property T, $p_1(f) = 0$ and so $f \xrightarrow{\perp} M$. The continuity of P_2 follows from the assumed continuity of $\phi''(f)(m, \cdot)$ and Lemma 2. Q.E.D.

Applying Theorem 1 with $n = 1$ and using the operators P_1 and P_2 of Theorem 3 gives two methods for determinig best approximations in M for $f \notin M$. Both may be viewed as descent methods for solving the system $\phi'(f - m)(m_i) = 0$, $i = 1, 2, \ldots, k$. The operator P_2 gives the damped Newton's method and generalizes the theorem of Karlovitz [6].

Iterating with the operator P_2, is in effect, the computation of a sequence of weighted L^2 approximations with varying weights. Since we assume for each $m \in M$ the map $X \backslash M \to X^*$ given by $h \to \phi''(h)(m, \cdot)$ is continuous the operator converges to

$$g \to \sum_{i=1}^{k} \phi''(f - P_M(f))(m_i, g)\, m_i$$

and if we choose the m_i orthonormal with respect to the inner product $\phi''(f - P_M(f))(\,,\cdot)$ we have a bounded projection with range M.

THEOREM 4. *Let $f \in X \backslash M$ and M have property T for ϕ. In addition assume that m_i, $1 \leqslant i \leqslant k$ is a basis for M orthonormal with respect to the inner product $\phi''(f - P_M(f))(\cdot, \cdot)$. If for each $m \in M$ the map $X \to X^*$ given by $g \to \phi''(g)(m, \cdot)$ is continuous in a neighborhood of $f - P_M(f)$, then $P_M'(f)\colon X \to M$ exists and is given by*

$$P_M'(f)(g) = \sum_{i=1}^{k} \phi''(f - P_M(f))(m_i, g)\, m_i\,.$$

Proof. Consider the function $Q: (X\backslash M) \times M \to M$ given by

$$Q(f, m) = \sum_{i=1}^{k} \phi'(f - m)(m_i)\, m_i .$$

The operator Q is defined and has continuous partial derivatives in a neighborhood of $(f, P_M(f))$ and $Q(f, m) = 0$ iff $m = P_M(f)$. A simple calculation shows that $Q_m'(f, P_M(f))(m) = -m$ and

$$Q_f'(f, P_M(f))(g) = \sum_{i=1}^{k} \phi''(f - P_M(f))(m_i, g)\, m_i .$$

From the implicit function theorem [5, p. 690], $P_M'(f)$ exists and

$$P_M'(f)(g) = -Q_m'(f, P_M(f))^{-1} (Q_f'(f, P_M(f))(g)).$$

Q.E.D.

Remark. The previous result can also be obtained by taking $\{m_1, \ldots, m_k\}$ and arbitrary basis for M and by considering the function from R^{k+1} to R given by

$$F(\lambda_1, \ldots, \lambda_k, \alpha) = \left\| f + \alpha m - \sum_{i=1}^{k} \lambda_i m_i \right\| .$$

The proof also uses the second derivative of the norm with the implicit function theorem and is given in Holmes and Kripke [3].

Note that since the argument of ϕ' and ϕ'' converges to $f - P_M f$ accelerating this convergence should improve the performance of the over all method.

Given any basis $\{m_1, \ldots, m_k\}$ for M we can define an "alternation" operator $\mathcal{A}: X\backslash M \to X$ by $\mathcal{A}(f) = g_k$ where $g_0 = f$ and $g_i = g_{i-1} - \lambda_i m_i$ with λ_i chosen to minimize $\| g_{i-1} - \lambda_i m_i \|$. This results in a "mixed" operator $P: X\backslash M \to M$ given by $P(f) = f - \mathcal{A}(f) + \alpha p(f)$ where

$$p(f) = \sum_{i=1}^{k} \phi'(\mathcal{A}(f))(m_i)\, m_i$$

and α is chosen to minimize $\| \mathcal{A}(f) - \alpha p(f) \|$. We again have that $\| f - P(f) \| < \| f \|$ unless $f \xrightarrow{\perp} M$ which implies convergence. It is obvious how to define the mixed operator in case M has property T.

5. Applications to L^p

If $1 < p < \infty$ then real L^p is uniformly smooth and uniformly rotund so that the hypotheses of Theorem 1 apply. Minimizing $\| f - m \|$ for $f \in X\backslash M$

is equivalent to minimizing $\phi(f - m) \equiv \| f - m \|^p$ and it is well known that ϕ has a Fréchet derivative $\phi'(f - m)(g) = p \int | f - m |^{p-1} \operatorname{sgn}(f - m)\, g\, d\mu$, i.e., $\phi'(f - m) = p \| f - m \|^{p-1} n(f - m)$. If $p \geqslant 2$ then ϕ' also has a Fréchet derivative with $\phi''(f - m)(g, h) = p(p - 1) \int | f - m |^{p-2} gh\, d\mu$, and for g fixed the map $L^p \to L^q$ given by $f - m \to | f - m |^{p-2} g$ is continuous in a neighborhood of $f - P_M(f)$. This is proved using geometric arguments in [3] and from a measure theory approach in [1].

If M is finite-dimensional and nonzero functions have support almost everywhere then property T follows from Hölder's inequality and compactness. Other conditions giving property T are discussed by Karlovitz [6] and Holmes and Kripke [3].

Numerical experiments with polynomial subspaces of $L^p[0, 1]$ seem to indicate that using the operator P_1 requires approximately the square of the number of iterations as when P_2 is used. However using P_2 involves solution of linear equations at each step and the matrices become illconditioned near the solutions, especially for large values of p. Calculating $\mathcal{O}\!\mathcal{l}(f)$ at each step (i.e., the "mixed" method for operator P_2) improves the situation since it accelerates convergence of the argument of ϕ'' and in some cases we can take $\alpha_2 = 1$ throughout and still have convergence. As the dimension of M increases it is best to break it into linearly independent subspaces and apply Theorem 1 in the general form. This idea can also be used to write M as $M_1 \oplus M_2$ where M_2 has property T and then alternate between P_1 and P_2.

References

1. B. Atlestam, Approximation Methods for Finite Dimensional Subspaces of L^p. Rapport 72.07, Insitutionen för Informationsbehandling, Chalmers tekniska högskola och Göteborgs Universitet.
2. R. Holmes, "A Course on Optimization and Best Approximation," Springer-Verlag, Berlin, 1972.
3. R. Holmes and B. Kripke, Smoothness of approximation, *Michigan Math. J.* **15** (1968), 225–248.
4. R. C. James, Orthogonality and linear functionals in normed linear spaces, *Trans. Amer. Math. Soc.* **61** (1947), 265–292.
5. L. V. Kantorovich and G. P. Akilov, "Functional Analysis in Normed Spaces," Pergamon Press, Elmsford, NY, 1964.
6. L. A. Karlovitz, Contruction of nearest points in the L^p, p even and L^∞ norms, *J. Approximation Theory* **3** (1970), 123–127.
7. W. J. Stiles, A solution to Hirshfeld's problem, *Nieuw. Arch. Wisk.* **13** (1965), 116–119.